Bioversity International

Bioversity International
Musa Germplasm
Transit Centre (ITC)

Factsheet 2

Unravelling the genetic basis of banana traits using Next Generation Sequencing technologies

For more information about the Bioversity International *Musa* Germplasm Transit Centre, see:

- Factsheet 1: *Operations, Outreach and Impact*
- Factsheet 3: *Phenotyping for drought*
- Factsheet 4: *Screening for resistance to Fusarium wilt*
- Factsheet 5: *Use of banana diversity for nutritious diets*

The availability of the banana reference genome sequence (D'Hont et al., 2012) and Next Generation Sequencing technologies have boosted research that helps to better understand the banana (*Musa* spp.) genetic diversity held in the Bioversity International *Musa* Germplasm Transit Centre (ITC). New insights into the genetics of the crop, associated with phenotyping, enable the exploration of banana genetic resources for a range of important end-users' traits, to support further uses to enhance food and nutritional security.

We are taking advantage of the latest Next Generation Sequencing technologies to i) produce high-density genetic markers for a large number of samples at low cost, and ii) to sequence, assemble and annotate key genomes, thus generating an unprecedented volume of information for researchers and breeders. Using the deluge of genomics data creates both a challenge and an opportunity. It places a unique challenge upon bio-informaticians to deal with the volume, speed and variety of data generated and to provide adequate storage, analyses and graphic visualization. At the same time, it provides a unique opportunity to unlock the potential of the genetic diversity stored in genebanks.

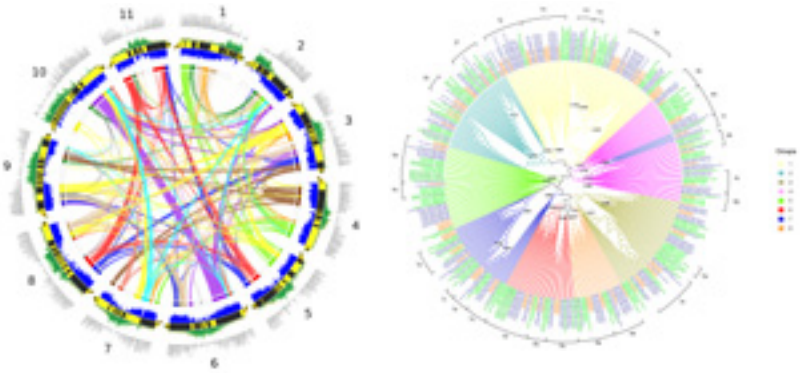

Figure 1: a) Circle representation of the banana chromosomes with their associated gene density and ancestral links. b) Evolutionary relationships between banana, rice, grape and Arabidopsis genes involved in stress responses such as drought. (Illustrations: Cirad/G. Martin & Bioversity/M. Rouard)

Top photo: Banana diversity at the germplasm collection in Trichy, India
Credit: Bioversity International/J. Sardos, edited by V. Guignon

Progress

- A high-quality reference genome sequence including annotated genes is available
- One third of the ITC collection is analyzed by high throughput genotyping
- Genome sequencing of key wild species and light sequencing of selected African cultivars (e.g. plantains) is underway
- Generated data are freely available through scientific databases for genetic and genomic resources
- Bioinformatics workflows are implemented for large-scale molecular markers and gene discovery
- Candidate genes for several traits including domestication and water use efficiency have been identified

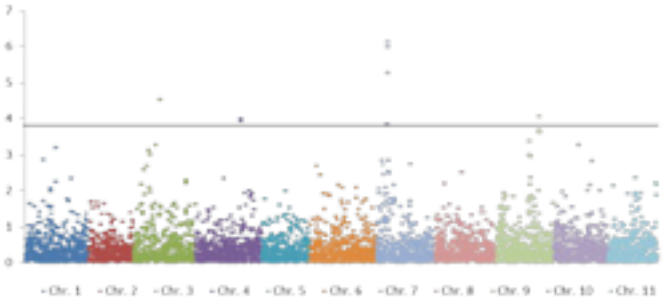

Figure 3: This chart, also known as a 'Manhattan plot', shows association mapping of a panel of ITC accessions selected for a Genome Wide Association study. Different colors depict banana chromosomes, and each square indicates a molecular marker whose position in the genome is known. The squares located above a certain threshold (dotted lines) are candidate regions in the genome correlated with traits related to domestication.

Figure 2: Participants at a 5-day training course in November 2013, Montpellier, France are learning about the use of *Musa* genome resources and bioinformatics tools. The training course was organized with the partners of the South, members of the Global *Musa* Genomics Consortium (GMGC), coordinated by Bioversity.

Looking ahead

Sequencing, genotyping and bioinformatics analyses are now integral parts of the research and management of genetic resources. They play an increasingly important role in understanding and exploiting the genetic variation of the crop's diversity, whether maintained in genebanks, on farm or in natural habitats, and as such contribute to enhance the productivity, sustainability and resilience of banana cultivars and their agricultural systems.

Selected publications:

1. D'Hont A, Denoeud F, Aury JM., Baurens FC, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M, et al. (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature, 10.1038/nature11241.
2. Droc, G., Lariviere, D., Guignon, V., Yahiaoui, N., et al. The Banana Genome Hub. Database (2013) doi:10.1093/database/bat035
3. Guignon V, Droc G, et al (2012) Chado Controller: advanced annotation management with a community annotation system. Bioinformatics. doi:10.1093/bioinformatics/bts046
4. Rouard M, Guignon V, et al. (2011) GreenPhylDB v2.0: comparative and functional genomics in plants. Nucleic Acids Res 39: D1095–1102. doi:10.1093/nar/gkq811.
5. Cenci A, Guignon V, Roux N, Rouard M (2014) Genomic analysis of NAC transcription factors in banana (*Musa acuminata*) and definition of NAC orthologous groups for monocots and dicots. Plant Mol Biol. doi:10.1007/s11103-013-0169-2.

Scientific databases:

1. *Musa* Germplasm Information System - www.crop-diversity.org/mgis
2. Banana Genome Hub - banana-genome.cirad.fr
3. GreenPhyl - www.greenphyl.org

Partners